

Improving the knowledge of the agritourist sector through integration between survey and administrative data

Roberto Gismondi¹, Maria Grazia Magliocchi¹, Filippo Oropallo¹, Francesco G.Truglia¹

¹*Italian National Statistical Institute, Italy*

Abstract

In Italy, agritourism is a particular secondary activity carried out by about 25 thousand agricultural holdings. Between 2010 and 2020, the number of agritourist farms increased of about 20%, while the overall number of farms decreased more than 30%. ISTAT (Italian National Statistical Institute) carries out a yearly survey on agritourism, which collects some structural data as agricultural surface, number of beds, number of place settings, services to customers. Available data do not include any economic indicators, as incomes, costs, investments, added valued and profits, as well as any information on employment. The estimation of agritourisms' economic results was possible through the integration between the survey microdata with administrative microdata derived from various sources. The ISTAT Farm Register (FR) contains data on the main economic activity, the technical-economic orientation, the standard output, the main crops cultivated, livestock, the size of the farms and the location of lands. Moreover, economic and employment indicators have been linked to the FR at the microdata level, through integration of these administrative sources: national social security, declarations relating to self-employed agricultural workers and agricultural labour, tax declarations and VAT returns, foreign trade. This exercise is carried since 2018 and is updated to 2022. Based on data matching, analyses on agritourisms' economic performance were possible, at a detailed territorial level. In particular, main results showed how, on average, agritourisms have higher productivity compared to other agricultural holdings and play an important role as regards the reduction of the historical gap between Northern and Southern Italy, since they are widespread in every Italian region. Through this integrated analysis, users can be provided with much more data than those available according to the statistical survey only.

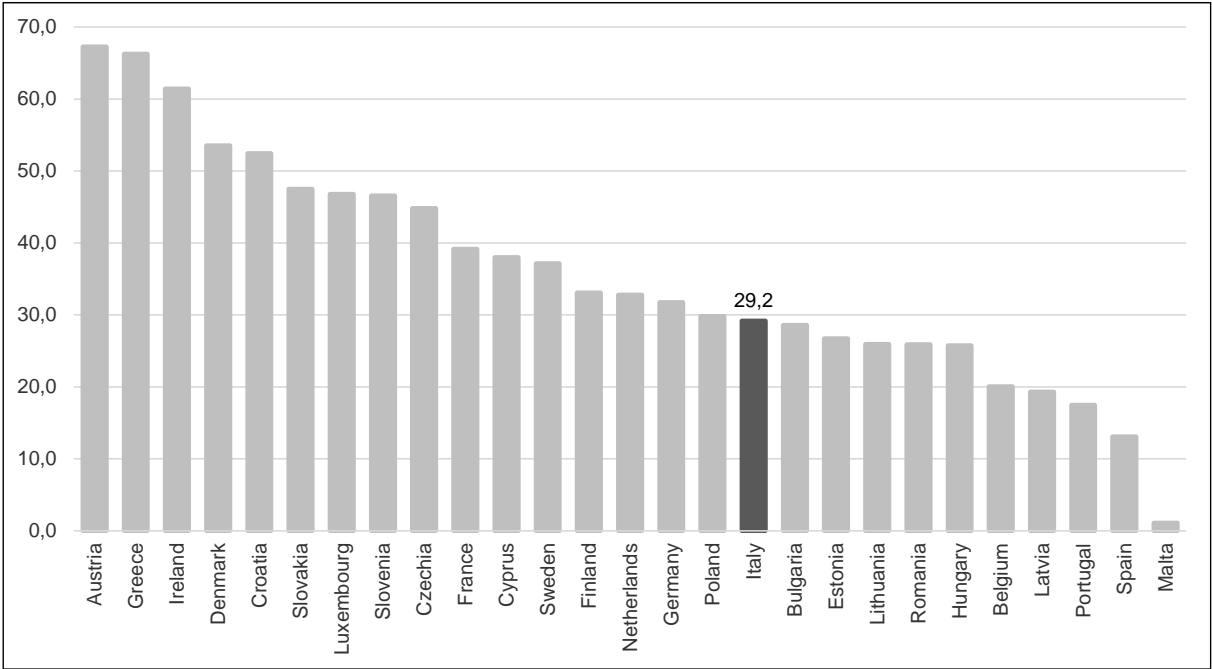
Keywords: Administrative data, Agriculture, Economic performance, Farm register, Matching

1. Why agritourism is important

Rural tourism, whose origins lie in agritourism and farm stays, is a type of tourism activity in which the visitor's experience is related to a wide range of products generally linked to nature-based activities, agriculture, rural lifestyle, culture, angling and sightseeing, providing visitors with a wide range of services, thus preventing depopulation. Starting from these considerations, the European Parliament (Šajn and Finer, 2023, 1) highlighted that it contributes to the economic, social and environmental sustainability of rural areas, in particular making a positive contribution to safeguarding small-scale and diverse farming, tackling social inequalities and creating employment opportunities for women. The EU Parliament has called on the European Commission to 'ring-fence' a specific allocation for agritourism under the new common agricultural policy. In Italy, in 2022 the percentage of nights spent in tourist

accommodation located in rural areas was 29,2% (Figure 1), while 32,7% is the share registered by the 19 Countries of Euro area. In the same year, in Italy arrivals in agriturismo facilities have exceeded four million: 3,4% of tourists choose agriturisms to spend their holidays. It must be kept in mind that agriturisms with overnight accommodation services make up just over 9,5% of the total agritourist structures. The ratio between Italian and foreign agritourists is 11 to 10. The nights spent exceed 15,5 million and of these 58% are due to foreign agritourists. The average length of stay, on average, is 3,8 days, 4,6 for foreigners and 3,1 for Italians. There is no doubt that these results are reflected in the economic performance of the sector: again in 2022, in fact, the current value of agriturismo production was equal to 1,5 billion euros, with an average annual growth compared to 2004 equal to +4,2%. In absolute terms, the production capacity of the sector has tripled in 8 years. For comparison, in the same period the average annual growth rate of the Italian agricultural sector was only 0,51% (ISTAT, 2024a, 5-6).

Figure 1: Share of nights spent in tourist accommodation in 2022 in rural areas of the EU



Source: Processing based on Eurostat data.

In this work, attention is paid to the analysis of the official statistical sources currently existing in Italy (section 2) as regards agriturisms. A comparison is proposed in section 3 that describes some potential causes of discrepancies among sources as regard the correct counting of the number of agritourist farms. The integration between the agriturisms survey microdata with administrative microdata derived from various sources led to the availability of additional indicators analysed in section 4, while section 5 analyses the odds ratios derived

from a logistic model, in order to identify the main reasons of discrepancies. Some prospective conclusions have been drawn in section 6.

2. Main data sources on agritourism in Italy

2.1 The ISTAT survey on agritourism

This yearly survey is carried out since 2004 and provides a detailed look on the evolution of agritourism in Italy. The data collection is based on administrative information concerning all farms authorized to carry out one or more types of agritourist activities (accommodation, catering, tasting and others). In Italy, agritourism is regulated by the Law 20 February 2006, n.96 which defines agritourism as reception and hospitality activities carried out by agricultural entrepreneurs. The information contained in the administrative archives comes mainly from the Municipalities, which issue authorizations for agritourist activities. Data are elaborated and transmitted to ISTAT by the Italian Regions, together with data on births and deaths of agritourisms. Data are released within one year from the reference year.

2.2 The 7th Census of Agriculture

The seventh general census of agriculture found its regulatory basis, at the European level, in Regulation (EU) 2018/1091¹. The agriculture census was mandatory, was carried out by all the European Union (EU) Member States and was referred to the date of 1 October 2020. Beyond the usual questions on agricultural surfaces, irrigation, livestock and labour force, the census collected data about the other gainful activities carried out by the farm beyond the basic agricultural production. Agritourism, care farming and educational farming were the additional farms' activities more oriented to the public measured by the census. They showed the highest growth respect to 2010 after energy production from renewable sources (ISTAT, 2024b).

2.3 The Extended Farm Register (EFR)

The EFR expands the information content of the standard Farm and as regards 2020 incorporated the Agriculture Census data through the integration of other data sources (Gismondi *et.al.*, 2021). The additional variables of the EFR are labor inputs (self-employer and employees) and their characteristics, labor costs, income statement variables such as sales and other revenues, changes in stocks, value of production, cost of goods and services, leasing cost, other charges. These variables allow us to compute the value added (VA) and the gross operating surplus (GOS) due to the agricultural activity for each farm. Other main

¹ <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32018R1091&from=FI>.

variables are capital cost, investments, import-export. The EFR is based on these data sources: 1) Statistical Business Registers and groups (for agricultural activities); 2) Social Security data (self-employed agricultural workers and agricultural employees); 3) Compliance survey by tax agency; 4) Tax return declarations (sole proprietorships, partnerships, corporates and non-commercial entities); 5) VAT returns; 6) Financial statements of corporate companies; 7) ISTAT foreign trade data; 8) Structural Business Statistics. The integration procedure was carried out using the ISTAT identifying code (SIM), which transforms the tax code of each unit or physical person into a unique anonymous code. The final estimate of the economic and structural variables of agricultural farms is based on data harmonization procedures, control of anomalous values and missing data imputation techniques.

3. Comparison among sources

As regards the year 2022, the matching process concerned the ISTAT survey on agritourisms and the Extended Farm Register (EFR) that incorporates the census results as well. Overall, the merge between the ISTAT surveys and the EFR led to the identification of 29.632 agritourisms, which in this framework represent the 100% of the whole theoretical reference population. The yearly survey counted 25.849 agritourist farms (the 87,2% with respect to the whole 29.632). Among them, 20.254 (the 68,3% of the whole) were counted as agritourisms in the EFR as well (Table 1). On the other hand, 5.595 units counted, as agritourisms by the survey (the 18, 9%) were not identified as agritourisms in the EFR; moreover, 3.783 units (the 12,8%) were identified as agritourisms in the EFR but not by the survey.

Overall, two basic reasons that can explain the lack of matching overlap. The first reason is the unavoidable discrepancy occurring when survey data and administrative data are compared. The survey respondents may declare a situation not consistent with the administrative process. Often, the date of reference of the administrative sources is not the same date of reference of data collected through the survey. Moreover, administrative sources may count as agritourisms farms that in practice did not start yet their agritourist activities, but that started the authorization phase only. The second reason is more concerned with problems related to the unit identifiers. Identifiers may be not complete or wrong; they may refer to different household persons than the farm manager (for instance, the wife of the farm holder); more than one farm may be associated with the same identifier. By the way, the Unique Identifier Project promoted by EUROSTAT may lead to improvements especially as regards the capability of national statistical systems to identify one and only one farm with the same identifier, no matter what the statistical or administrative source concerned could be.

Table 1: Comparison between the ISTAT survey on agritourisms and the EFR – Absolute and percent figures. The year 2022 (Yes = the farm is an agritourism)

	EFR	No	Yes	Total
Survey	No	0	3.783 (12,8)	3.783 (12,8)
	Yes	5.595 (18,9)	20.254 (68,3)	25.849 (87,2)
	Total	5.595 (18,9)	24.037 (81,1)	29.632 (100,0)

Source: processing based on ISTAT data: survey on Agritourisms and the EFR.

Among the 5.595 units counted as agritourisms by the survey but not identified as agritourisms in the EFR, 3.674 units were not matched with any other unit in the EFR (NoEFR) at all. This subset was matched through statistical matching procedures (Kim and Shao, 2014). The Table 2 shows the percent ratio between the number of NoEFR and the number of agritourisms identified by the EFR. Overall, about one unit on eight (12,4%) is not matched with any agritourisms in the EFR. The distribution by geographical area is quite heterogeneous: the weight of NoEFR is larger in the South (20,7%) and quite low in Northeast (6,5%).

Table 2: Units counted as agritourisms by the ISTAT survey not matched with any other unit in the ISTAT EFR – percentage values on the number of agritourisms

Geographic area	%
North west	12,3
North East	6,5
Centre	13,4
South	20,7
Islands	15,9
Italy	12,4

Source: processing based on ISTAT data: Survey on Agritourisms and the EFR.

4. Additional indicators derived from the EFR

The large set of indicators available through the EFR increases the capability of analysing farms' diversity due to multifunctionality (Van der Ploeg *et.al.*, 2009; Etumnu and Gray, 2020). For 2022 it is possible to classify the 25.849 agritourisms of the yearly survey by kind of statistical unit (Table 3). Moreover, we can analyse the economic indicators reported in the Table 4 (by size classes) and 5 (by geographic area). For instance, productivity (value added per worker) increases as size increases, with an average level of 32.600 euros. Agritourisms with 5 or more employees have profitability higher than average. Productivity halves in the

Southern regions, with negative gross operating margins (-37,1% in the South and -40,6% in the Islands), while in the North and the Centre productivity is above the national average.

Table 3: Number of agritourist farms (yearly survey) by kind of statistical unit. The year 2022

Kind of unit	Agritourist farms		Units in the EFR		Agritourist farms % share
	Absolute figures	% figures	Absolute figures	% figures	
Agricultural enterprises	14.551	56,3	397.624	33,3	3,7
Enterprises with secondary agricultural activity	8.758	33,9	63.563	5,3	13,8
Partially market active farms or non-profit	2.340	9,1	468.480	39,2	0,5
Farms mainly for self-consumption	200	0,8	265.966	22,2	0,1
Total	25.849	100,0	1.195.633	100,0	2,2

Source: processing based on ISTAT data: Survey on Agritourisms and the EFR.

Table 4: Economic indicators of agritourist farms by size classes. The year 2022

Employees	Units	Workers per unit	Turnover per worker (000€)	Value added per worker (000€)	Unit Labour Cost (000€)	Gross profitability %	Foreign turnover %
1	11.355	1,0	46,8	18,1	21,1	-16,6	1,0
2	6.149	2,0	39,8	22,6	22,0	2,8	2,2
3-4	5.929	3,5	50,4	27,6	23,7	14,3	4,6
5-9	1.460	6,8	66,3	35,7	24,9	30,2	8,5
10-19	635	12,8	90,7	41,1	25,9	36,9	12,5
20 and more	321	43,9	119,1	52,8	25,2	52,3	39,4
Total	25.849	3,0	67,2	32,6	24,7	24,3	17,0

Source: processing based on ISTAT data: Survey on Agritourisms and the EFR.

Table 5: Economic indicators of agritourism farms by geographical area (NUTS1). The year 2022

NUTS1	Units	Workers per unit	Turnover per worker (000€)	Value added per worker (000€)	Unit Labour Cost (000€)	Gross profitability %	Foreign turnover %
North west	3.956	3,1	73,6	38,4	26,1	32,0	8,6
North east	7.454	2,8	62,6	35,7	25,6	28,2	5,1
Centre	9.364	3,2	77,8	35,6	25,1	29,5	29,7
South	3.323	2,5	47,3	15,0	20,5	-37,1	1,9
Islands	1.752	2,9	41,3	16,3	22,9	-40,6	18,6
Italy	25.849	3,0	67,2	32,6	24,7	24,3	17,0

Source: processing based on ISTAT data: Survey on Agritourisms and the EFR.

We compared the EFR economic performance indicators referred to 2019 (before the pandemic) to the 2022 indicators. Agritourisms' productivity increased (+22,1%). The growth characterized each size class except the largest (20 and more employees, Table 6) .

Table 6: Economic indicators of agritourist farms by size classes. % changes between 2019 and 2022

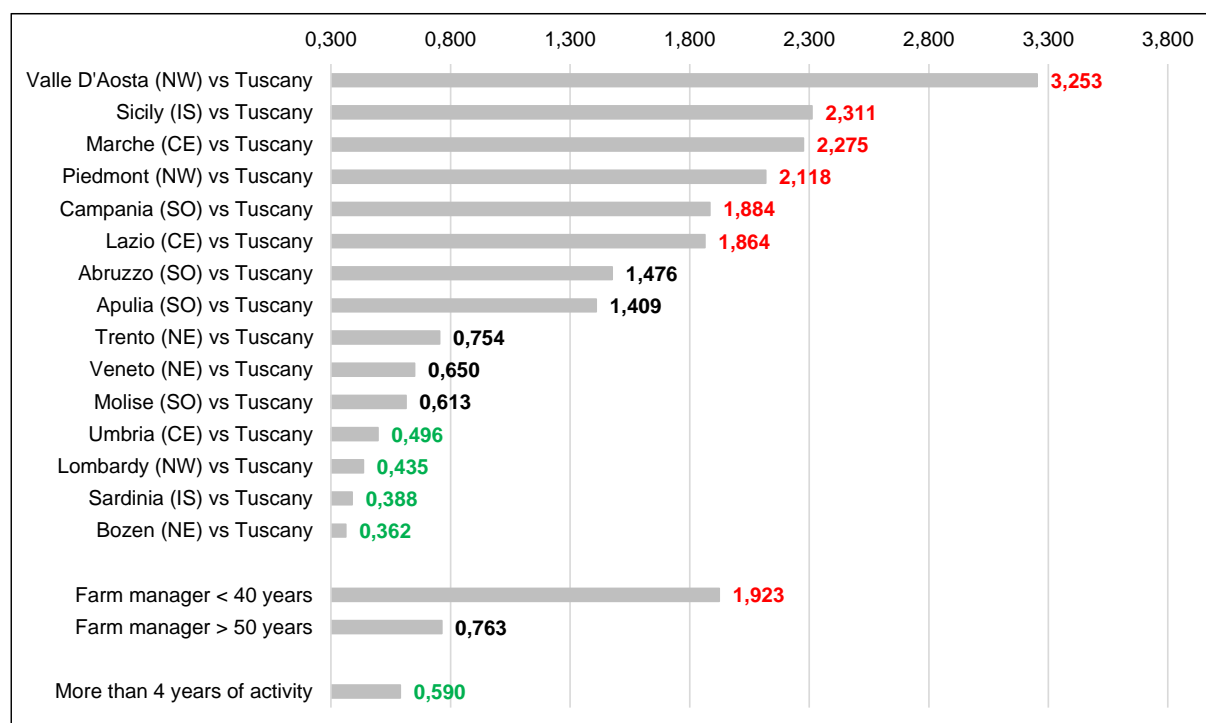
Employees	Units	Workers	Turnover per worker	Value added per worker	Unit Labour Cost (000€)	Gross profitability	Foreign turnover
1	1,0	-1,3	41,2	1,8	18,0	-	18,6
2	1,5	0,5	17,4	18,7	20,3	-31,2	65,1
3-4	10,6	0,7	21,6	21,8	21,3	3,0	50,9
5-9	22,9	-0,7	15,8	15,9	14,4	3,1	32,4
10-19	26,5	2,6	27,4	20,0	10,5	17,3	22,2
20 and more	42,0	0,7	-0,1	-4,0	-11,7	8,7	3,8
Total	5,2	9,8	22,1	15,9	6,5	37,7	11,8

Source: processing based on ISTAT data: Survey on Agritourisms and the EFR.

5. The logistic model

Logistic regression (Hilbe, 2009) is used for modeling the probability of an event (dependent variable Y) through a series of explanatory X -variables. In this context, the Y binary variable is equal to 1 if the farm is classified as agritourism by both the yearly survey and the EFR, and is equal to 0 otherwise. If p is the probability that $Y=1$, the ratio $p/(1-p)$ is the *odds ratio* and is equal to 1 if and only if $p=0,5$. The odds ratios more different from 1 identify the farms' features that influence more the probability that the farm is concordant. Figure 2 shows the main results.

Figure 2: Odds ratios derived from the logistic model for the most significant dependent variables



Source: processing based on ISTAT data: Survey on Agritourisms and the EFR. The regional odds ratios have been calculated using Tuscany (the region with the higher number of agritourisms) as benchmark.

The explicative factors with odds ratios quite higher than one concern features of the farm that increase the probability that the farm is present as agritourism in both sources; on the other hand, the factors with odds ratios quite lower than one concern characteristics that increase the probability that the farm is not classified as agritourism in both sources. The most important factor that increase the probability that the farm is an agritourism in both sources are: the location in 6 out of the 21 Italian Regions belonging to North, Centre, South or Islands (Valle D'Aosta, Sicily, Marche, Piedmont, Campania and Lazio); to have a “young” farm manager (less than 40 years). On the other hand, the most important factor that reduce the same probability are: location in 4 out of the 21 Italian Regions belonging to North, Centre and Islands (Umbria, Lombardy, Sardinia and Bozen); to practice agritourism since more than 4 years.

6. Perspective conclusions

Looking ahead, we must continue comparing sources on agritourisms and removing obstacles that prevent their convergence. At the same time, it is more and more important to access the administrative databases from a micro perspective. The informative value added through record linkage is huge and users may benefit from the additional statistical indicators available. The same exercise may be replied as regards other gainful activities carried out by farms beyond agritourist services, as production of energy from renewable sources for instance.

References

- Šajn, N., & Finer, K. (2023). Rural Tourism. Members' Research Service. PE751.464 – EU Parliament. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751464/EPRS_BRI\(2023\)751464_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751464/EPRS_BRI(2023)751464_EN.pdf).
- Etumnu, C., & Gray A.W. (2020). A Clustering Approach to Understanding Farmers' Success Strategies. *Journal of Agricultural and Applied Economics*. 52/2020. 335–351. [doi:10.1017/aae.2020.4](https://doi.org/10.1017/aae.2020.4).
- Gismondi, R., Magliocchi, M.G., Oropallo, F., & Truglia, F. (2020). Integration of farms microdata: economic analysis of agritourisms and impact assessment of the effects of COVID-19. *Rivista di statistica ufficiale*. 1/2021. 4-30. https://www.istat.it/it/files/2021/05/RSU-1_2021_Article-4.pdf.
- Hilbe J.M. (2009). *Logistic regression models*. Chapman & Hall. <https://www.routledge.com/Logistic-Regression-Models/Hilbe/p/book/9781138106710>.
- ISTAT (2024a). Le aziende agrituristiche in Italia. Anno 2022. <https://www.istat.it/it/archivio/292868>.
- ISTAT (2024b). Settimo censimento generale dell'agricoltura: dati definitivi per centro aziendale. <https://esploradati.istat.it/databrowser/#/it/censimentoagricoltura/categories/CENSAGR>.
- Kim, J.K., & Shao., J. (2014). *Statistical Methods for Handling Incomplete Data*. Boca Raton, FL, U.S.: Chapman and Hall/CRC Press.
- Van der Ploeg, J.D., Laurent, C., Blondeau, F., & Bonnafous, P. (2009). Farm diversity, classification schemes and multifunctionality. *Journal of Environmental Management*. 90/2009. 124-131. <https://doi.org/10.1016/j.jenvman.2008.11.022>.